

## 특집: 미생물계능 및 분석방법(III)

### 염색체 DNA의 새로운 분석 기법

#### - Natural host 내에서 염색체 DNA 특정부위의 선택적 대량 증폭법 -

윤영걸 · 김선창

한국과학기술원 생물학과

특정 염색체 DNA 일부를 cloning 하는 것과 그것의 염기서열을 밝혀내는 것은 한 생물체의 구조와 생물학적 기능을 이해할 수 있는 열쇠가 될 수 있다. 또한 염색체 DNA의 분자구조를 설명할 수 있는 궁극적인 방법 역시 그것의 염기서열을 정확히 기술하는 것이다. 현재까지, 거대한 염색체 DNA를 분석하는 방법은 다음 몇가지 단계로 구분할 수 있다.

- (1) 염색체의 유전자 지도작성 (genome mapping)
- (2) 염기서열 결정 (sequencing)을 위한 DNA 절편의 준비
  - a. 큰 DNA 절편의 준비
  - b. 염기서열 결정반응을 바로할 수 있도록 준비된 작은 DNA 절편의 확보
- (3) 염기서열 결정반응의 수행
- (4) 염기서열의 조합
- (5) 염기서열의 분석

이러한 분석법은 직접적으로 염기서열을 결정하는 (2)~(4) 단계에 어떻게 접근하는가에 따라 'bottom-up approach'와 'top-down approach'라는 두 가지의 기본적인 개념이 등장하였다(4).

#### Bottom-up Approach

Bottom-up approach는 Fred Sanger가  $\lambda$  phage의 genome을 sequencing할 때 선택한 방법(26)으로서, 먼저 sequencing할 염색체 DNA 절편을 random하게 선택한 다음, 역시 random하게 sequencing 반응을 수행하는 것이다. 얻어진 모든 sequence data는 computer를 이용하여 sequence가 겹쳐지는 부분을 연결하여 하나의 완전한 sequence를 만든다. 이러한 접근법을 이용한  $\lambda$  phage genome 분석이 성공할 수 있었던 이유는  $\lambda$  DNA의 크기가 비교적 작았으며 매우 안정된 clone을 확보할 수 있었기 때문이었다.

그러나, Prokaryotic, eukaryotic DNA 절편들은 cloning host에서 발현될 수 있는 수많은 gene들 혹은 pseudogene들을 가지고 있으며 또한 반복적인 sequence 등의 다양한 형태적인 특징들을 가지고 있다. 이런 요인들은 cloning host에 대해 독성을 유발하거나 그 clone에 대하여 불안정성을 나타나게 하

며 결과적으로 toxic하거나 unstable한 부분의 DNA를 deletion하거나 rearrangement를 일으키는 강한 선택력을 갖게 만든다. Bottom-up approach는 외래 DNA를 cloning host에서 다루게 되므로 host와 DNA 사이에 새로운 balance가 유발되며 이는 곧 하나의 "evolution"을 의미한다. 즉, 이것은 cloning된 DNA가 원래 있었던 자신의 sequence 보다는 cloning host에서 유지될 수 있는 형태로의 변형을 뜻하며 이와같은 현상으로 인해 eukaryotic genome 분석이 매우 어려워지게 된다. 또한 완전한 genome분석을 위해서는 충분한 genomic clone을 보유하고 있어야 하며, toxic segment의 존재로 말미암아 많은 절편들을 cloning할 수 없다면 결국 그 부분의 sequence gap이 나타나게 된다. 이외에도 각 clone들 사이에서 중복되는 부분을 일일이 찾아서 하나의 연속되는 sequence를 작성해야하므로 bottom-up approach는 거대한 genome 전체를 분석하는데 있어서 선결되어야 하는 많은 문제점들을 내재하고 있는 것이다.

#### Top-down approach

Top-down approach는 한 세포의 genome으로부터 직접 염색체 DNA 절편을 분리하는 방법으로서, random cloning 방법에서 사용하는 외래 vector/host system을 요구하지 않는다. 이 방법은 비교적 완전한 map을 얻을 수 있는 접근법으로서 그 원리는 먼저 genome의 유전자 지도에 근거하여 일련의 strain을 조작한 다음, sequencing할 organism의 genome을 순서적으로 그 세포로부터 직접 분리하는 것이다. 이 접근법에서는 physical mapping을 하는 것과 염색체 DNA 절편을 분리하는 것이 서로 밀접하게 연관되어 있다. 왜냐하면, 이 접근법에서 제시된 방법이 서로 인접하는 염색체 DNA 절편을 각각 분리하는 것으로서 궁극적으로는 전체 genome을 총 망라할 수 있기 때문이다. 이미 physical map이 있는 경우 이러한 접근법은 다음과 같은 방법에 응용할 수가 있다.

**in vitro enzymatic excision: 50~100 kb 크기의 단일 copy 절편의 생성**

분석하고자하는 genome이 일단 100 kb 정도의 단위로 유전자 지도가 작성되어 있다면(7), *NotI* 등의 rare-cutting enzyme를 사용하여 genome을 절단하거나 혹은 어떤 sequence-tagged site (STS)를 이용한 RecA-mediated Achilles' heel (AC) cleavage 방법(16,17,18)을 병행하여 원하는 부분만을 50~100 kb의 크기로 절단한 다음, 이들을 pulsed-field gel electrophoresis (PFGE) 혹은 affinity-capture method 등을 사용하여 분리할 수가 있다. 이 접근법은 어떠한 cloning 단계도 거치지 않고 직접 sequencing할 genome으로부터 DNA를 얻을 수 있는 장점이 있으나, genome을 절단하기 위한 순수분리된 enzyme이 있어야 하며 또한 genome 하나당 원하는 부분 하나만을 얻을 수 있기 때문에 그 efficiency는 상당히 낮다.

**in vivo excision 및 이와 연관된 amplification: 특이적인 excision system을 이용한 단일 copy 절편의 생성 및 이의 증폭**

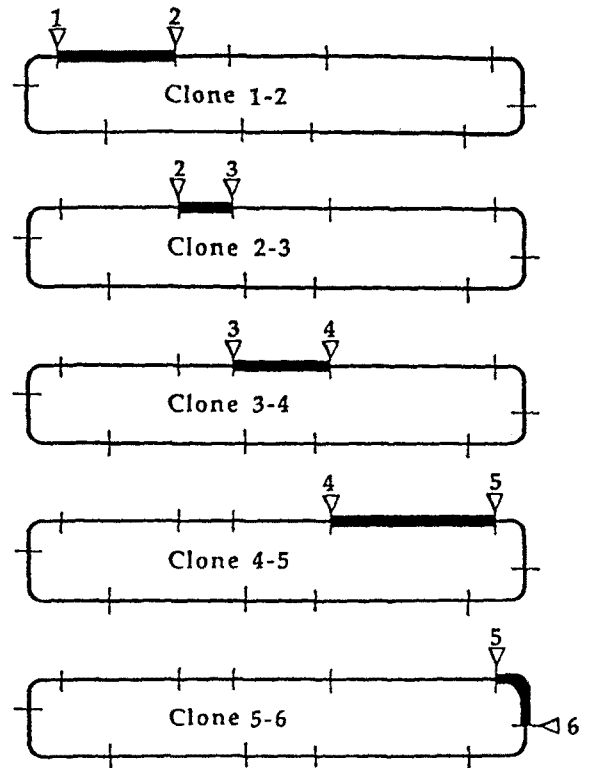
(1) *In vivo* excision

이것은 *in vitro* excision에서 요구되는 순수분리된 enzyme이 필요하지 않으며 또한 어떤 amplification system과 결합을 시킬 수 있다면, genome 하나당 하나의 DNA 절편만을 얻는 낮은 efficiency도 극복할 수 있다. 여기에서 사용하는 excision system은 site-specific recombination(1,2,20)으로서 이것은 그 특성상 recombinase가 인식하는 인지부위 (recognition site) 사이의 DNA를 inversion 시키거나 deletion 되어 나오게 한다. 예를 들어, 염색체 DNA 상에 recombination target site 두 개가 같은 방향으로 일정거리 떨어져서 놓여 있을 때 recombination이 일어나면 두 개의 인지부위 사이의 DNA는 원형의 DNA로서 염색체 DNA에서 떨어져 나오게 된다. 이러한 특성을 이용하면 염색체 DNA의 특정부위 사이의 DNA를 고전적인 cloning 방법을 동원하지 않고서도 쉽게 plasmid의 형태로 바로 자신의 세포 내에서 cloning할 수가 있는 것이다. 이러한 system에는 phage λ에서 유래한 att/Int system (20), yeast 2μ plasmid의 FRT/FLP system (10,15,34), phage P1의 lox/Cre system (12,13,30) 등이 있으며 이것을 어떤 virus나 plasmid의 특이적인 복제기구와 결합시키면 자체세포의 genome에서 떨어져나온 염색체 DNA 절편이 자체세포 내에서 복제, 증폭할 수가 있게 된다. 이러한 접근법의 기본 개념은 λ prophage가 host genome에 삽입되어 있다가 어떤 신호에 의해 정확하게 excision 되어 50 kb 정도되는 자신의 DNA를 amplification하는 현상에서 착안되었다.

.) Placing the sites on the genome to be sequenced

위의 개념을 실제화하기 위해서는 먼저 recombinase의 인지부위 (excision-mediating sites, ems; att, FRT, lox)를 세포의 viability 혹은 genome의 stability에는 영향을 주지 않는 위치에 약 50~100 kb 정도의 간격으로 삽입시켜야한다. Random한 삽입 방법으로는 Singer등(29)이 *E. coli* library를 만들때

사용한 *Tn10*을 이용하는 것으로 transposons (bacteria에 대해서는 *Tn5*, *Tn10*; Drosophila에 대해서는 P element (3,8)과 ems, 그리고 selection marker를 하나의 작은 plasmid에 실은 다음, 이것을 이용하여 다량의 marked strain을 만든다. 이 방법을 사용할 때는 반드시 PFGE나 이와 연계한 제한 효소적 방법을 적용하여 mapping을 해야 한다. 또 하나의 방법은 site-specific insertion으로서, Rec system, 즉 homologous recombination(19)을 이용하는 것이다. 하나의 작은 plasmid에 insertion하고자하는 위치와 homologous한 DNA 절편(예, lac, gal, trp), ems, 그리고 selection marker를 실은 후, 이것을 세포 내로 주입시켜 homologous recombination을 유도하여 genome에 삽입시킬 수 있다. 이와같은 방법을 이용하면 원하는 위치에 하나의 ems를 가진 각각의 library strain을 구축할 수 있으며, ems가 50~100 kb 간격으로 떨어져 있는 두 strain을 유전학적인 방법으로 crossing 함으로써 두 개의 ems가 같은 genome 상에 50~100 kb 떨어져 위치할 수 있게 만들 수 있다. 약 100 kb 단위로 genome을 나누고자할 때 *E. coli*의 경



**Fig. 1.** Saturation of the bacterial circular genome with the ems-flanked segments. The bold-line fragments are consecutively arranged, because the same ems element was used for crossing with neighboring ems elements to the left and to the right. For *E. coli*, only, 50 strains are required to prepare 50 consecutive approx. 100-kb fragments which would completely cover the entire *E. coli* genome.

우에는 약 50개의 strain이면 *E. coli* genome 전체를 포함할 수 있으며 *Drosophila*의 경우는 약 1500개 strain이면 전체 genome을 완전히 확보할 수 있다. 이 접근법의 장점은 한 strain에서 원하는 부분의 염색체 DNA 절편을 얻고 곧이어 다음 strain에서 그 다음의 염색체 DNA 절편을 연속적으로 얻을 수 있다(그림 1)는 것과 자신의 세포내에서 직접 sequencing할 DNA를 분리하는 것으로 random cloning 방법에서 제기되는 여러 문제점들을 피해 나갈 수가 있다.

## 2) The lox/Cre system

이 system(1,2)은 bacteriophage P1에서 유래한 것으로 site-specific recombinase Cre는 자신의 target site인 lox site 두 개를 인지하여 inversion 혹은 excision 등의 recombination을 일으킨다. lox site는 13 bp의 inverted repeat와 8 bp의 spacer로 이루어진 13+8+13 bp 구조를 가지고 있다. lox/Cre system은 *in vivo* 혹은 *in vitro*에서도 작용하며 bacteria는 물론 plant, mammalian 등의 eukaryote에서도 효과적으로 적용할 수 있음이 보고되어 있다(12,27). 이 system은 다른 cofactor는 요구되지 않고 단지 Cre protein 만이 필요하며 적당한 host-specific ori를 삽입하게 되면 *in vivo* excision에 뒤따르는 amplification을 촉진할 수 있다.

## 3) The FRT/FLP system

이 system은 yeast의 2 $\mu$  plasmid에서 유래한 것으로 Cre/lox system과 매우 유사하다(34). FRT site는 8 bp의 spacer를 가진 13 bp inverted repeat를 구조를 가지나 그 sequence는 Cre/lox system과는 다르며 이 system도 역시 eukaryote에서 효과적으로 적용할 수 있다고 알려져 있다(20). 이러한 excision system에서 중요한 것은 host의 viability나 stability에 영향을 주지않도록 매우 stringent한 control system이 필요하다는 것이다.

## 4) The att/Int system

이것은 bacteriophage 에서 유래한 system으로서 두 개의 attL site 혹은 attP와 attB site를 이용할 수 있다. 이 element를 사용할 때 필요한 것은 excision을 위한 integrase이고 IHF (integration host factor)는 host에서 직접 공급하게된다. 이 system의 가장 큰 장점은 recombination 반응이 비가역적으로 일어나므로 일단 excision이 일어난 후에는 re-insertion 되지 않는다는 점이다.

## (2) Amplification system

Replication system을 excision element와 함께 genome 상에 도입하게되면 *in vivo* excision된 염색체 DNA 절편을 excision과 동시에 증폭시킬 수가 있다. 즉, site-specific recombination과 ori-specific replication system을 병행하게되면, 원하는 부분의 DNA는 site-specific recombination에 의해 genome에서 원형의 DNA로 떨어지거나오게되며 이 원형의 생물산업

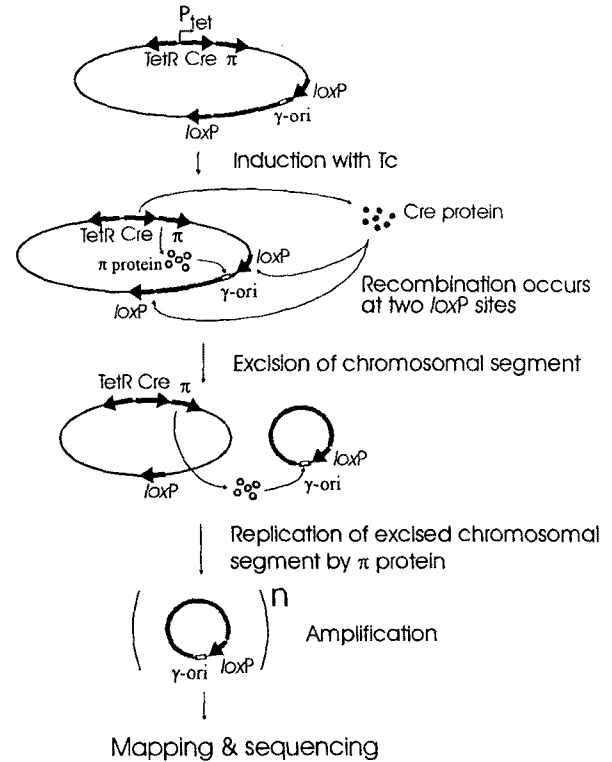


Fig. 2. A schematic diagram of Cre/loxP-mediated excision and  $\pi/\gamma$ -ori mediated amplification of chromosomal DNA.

DNA는 ori-specific replication에 의해 증폭할 수 있게 되는 것이다(그림 2). 이러한 construct는 다양한 prophage나 provirus에서 흔히 나타나는 현상으로서 적당한 excision/replication system을 조합할 경우, genome의 원하는 부분만을 특이적으로 분리, 증폭할 수 있으며 genome sequencing에 필요한 만큼의 DNA를 충분히 확보할 수 있다. 모든 amplification system도 역시 매우 tight 하게 조절할 수 있어야 한다. 따라서 어떤 induction signal에 의하여 excision이 일어남과 동시에 replication에 의한 amplification을 유도하는 것이 유리하다. 이러한 control system으로 시도되고 있는 것으로 두 가지 type이 있다.

(a) 어떤 replication protein의 합성을 조절하는 방법으로  $\lambda$  phage의 O, P protein, R6K plasmid의  $\pi$  protein (31,32), 2 $\mu$  plasmid의 Flp protein(6) 등이 있다.

(b) inducible promoter를 이용하여 ori element와 replication primer의 조절하는 방법으로서,  $\lambda$  system에 있어서는 ori transcription의 조절, 또는 plasmid 복제기구에 있어서는 RNA primer 합성을 조절(11,37,39) 하는 방법 등이 있다.

잘 알려져 있는 replication origin으로는 다음 몇 가지를 들 수 있다. gram bacteria에 대해서는 broad-host-range를 가진 RK2 ori (9), yeast 2 $\mu$  plasmid에서 유래한 ARS element(6), mammalian 세포에 대해서는 Epstein-Barr virus(EBV)의

EBNA-1 protein과 *oriP* element (25,38) 등이 있다.

## 결 론

염색체 DNA는 일반적으로 그 크기가 매우 크므로 이를 분석하여 각 유전자의 상대적인 위치와 유전자들의 배열순서 및 그들의 일차원적인 염기서열을 결정하는 것은 매우 중요하면서도 상당한 시간과 비용을 요하는 어렵고 복잡한 일로서 인식되어 왔다. 그러나, 이 글에서 제시한 접근법 (Top-down approach)을 이용한다면 전체 genome을 포함할 수 있는 연속적인 100 kb 절편을 자체세포에서 직접 분리할 수 있으며 또한 분리한 DNA를 sequencing 반응의 template로서 사용할 수 있게 되므로 염색체 DNA의 분석을 아주 쉽고 빠르게 할 수 있으며 또한, 고전적 cloning 방법에서 나타나는 infidelity 문제를 직접 해결할 수 있으므로 염색체 DNA의 구조분석에 매우 강력한 도구로 사용할 수 있을 것이다.

## REFERENCES

1. Abremski, K., Hoess, R. and Sternberg, N. 1983. Studies on the properties of P1 site-specific recombination: Evidence for topologically unlinked products following recombination. *Cell* **32**, 1301-1311.
2. Abremski, K., and Hoess, R. 1984. Bacteriophage P1 site-specific recombination. Purification and properties of the Cre recombinase protein. *J. Biol. Chem.* **259**, 1509-1514.
3. Benz, W. K. 1989. Transposition and Excision of P Elements in the Genome of *Drosophila melanogaster*, and the Localization of Repressor-producing Elements. Ph.D. Thesis, University of Wisconsin, Madison, WI.
4. Billings, P. R., Smith, C. L. and Cantor, C. R. 1991. New techniques for physical mapping of the human genome. *FASEB. J.* **5**, 28-34.
5. Botstein, D., Falco, S. C., Stewart, S. E., Brennan, M., Scherer, S., Stinchcomb, D. T., Struhl, K. and Davis, R. W. 1979. Sterile host yeasts (SHY): A eukaryotic system of biological containment for recombinant DNA experiments. *Gene* **8**, 17-24.
5. Brewer, B. J. and Fangman, W. L. 1987. The localization of replication origins on ARS plasmids in *S. cerevisiae*. *Cell* **51**, 463-471.
7. Davies, K. E. and Tilgham, S. M. 1990. Genome analysis in genetic and physical mapping. Vol. **1**, 7-8.
8. Engels, W. R. P. 1989. Elements in *Drosophila*. In: Berg, D. and Howe, M.(Eds), *Mobil DNA*, American Society of Microbiology, Washington, DC, pp437-484.
9. Fang, F. C., Durland, R. H. and Helinski, D. R. 1993. Mutations in the gene encoding the replication-initiation protein of plasmid RK2 produce elevated copy numbers of RK2 derivatives in *Escherichia coli* and distantly related bacteria. *Gene* **133**, 1-8.
10. Gates, C. A. and Cox, M. M. 1988. FLP recombinase is an enzyme. *Proc. Natl. Acad. Sci.* **85**, 4628-4632.
11. Gil, D. and Bouche, J.-P. 1991. ColE1-type vectors with fully repressible replication. *Gene* **105**, 17-22.
12. Hoess, R. H. and Abremski, K. 1984. Interaction of the bacteriophage P1 recombinase Cre with the recombining site *loxP*. *Proc. Natl. Acad. Sci.* **81**, 1026-1029.
13. Hoess, R. H. and Abremski, K. 1985. Mechanism of strand cleavage and exchange in the Cre-*lox* site-specific system. *J. Mol. Biol.* **181**, 351-362.
14. Hoess, R. H., Wierzbicki, A. and Abremski, K. 1986. The role of the *loxP* space region in P1 site-specific recombination. *Nucleic Acids Res.* **14**, 2287-2300.
15. Huang, L.-C., Wood, E. A. and Cox, M. M. 1991. A bacterial model system for chromosomal targeting. *Nucleic Acid Res.* **19**, 443-448.
16. Koob, M., Grimes, E. and Szybalski, W. 1988. Conferring operator specificity on restriction endonucleases. *Science* **241**, 1084-1086.
17. Koob, M. and Szybalski, W. 1990. Cleaving yeast and *Escherichia coli* genomes at a single site. *Science* **250**, 271-273.
18. Koob, M., Burkiewicz, A., Kur, J. and Szybalski, W. 1992. RecA-AC single-site cleavage at any predetermined restriction site. *Nucleic Acid Res.* **20**, 5831-5836.
19. Kowalczykowski, S. C., Dixon, D. A., Eggleston, A. K., Lauder, S. D. and Rehrauer W. M. 1994. Biochemistry of homologous recombination in *Escherichia coli*. *Microbiol. Rev.* **58**, 401-465.
20. Landy A. 1993. Mechanistic and structural complexity in the site-specific recombination pathways of Int and FLP. *Curr. Opin. Genet. Develop.* **3**, 699-707.
21. Link, A. J. and Olson, M. V. 1991. Physical map of the *Saccharomyces cerevisiae* genome at 110kb resolution. *Genetics* **127**, 681-698.
22. Mack, A., Sauer, B., Abremski, K. and Hoess, R. 1992. Stoichiometry of the Cre recombinase bound to the *lox* recombining site. *Nucleic Acids Res.* **20**, 4451-4455.
23. Meselson, M. and Stahl, F. W. 1958. The replication of DNA in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **44**, 671-682.
24. Posfai, G., Koob, M., Hradecna, Z., Hasan, N., and Filutowicz, M. 1994. In vivo excision and amplification of large segments of the *Escherichia coli* genome. *Nucleic Acid Res.* **22**, 2392-2398.
25. Reisman, D. and Sugden, B. 1986. trans Activation of an Epstein-Barr viral transcriptional enhancer by the Epstein-Barr viral nuclear antigen 1. *Mol. Cell. Biol.* **6**, 3838-3846.
26. Sanger, F., Coulson, A. R., Hong, G. F., Hill, D. F. and Petersen, G. B. 1982. Nucleotide sequence of bacteriophage  $\lambda$

- DNA. *J. Mol. Biol.* **162**, 729.
27. Sauer B. and Henderson, N. 1989. Cre-stimulated recombination at *loxP*-containing DNA sequences placed into the mammalian genome. *Nucleic Acid Res.* **17**, 147-161.
  28. Sikorski, R. S. and Hieter, P. 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122**, 19-27.
  29. Singer, M., Baker, T. A., Schnitzler, G., Deischel, S. M., Goel, M., Dove, W., Jaacks, K. J., Grossman, A. D., Erickson, J. W. and Gross, C. A. 1989. A collection of strains containing genetically linked alternating antibiotic resistance elements for genetic mapping of *Escherichia coli*. *Microbiol. Rev.* **53**, 1-24.
  30. Sternberg, N., and Hamilton, D. 1981. Bacteriophage P1 site-specific recombination. I. Recombination between *loxP* sites. *J. Mol. Biol.* **150**, 467-486.
  31. Stalker, D. M., Kolter, K. and Helinski, D. R. 1982. Plasmid R6K DNA replication I. Complete nucleotide sequence of an autonomously replicating segment. *J. Mol. Biol.* **161**, 33-43.
  32. Kolter, K. and Helinski, D. R. 1982. Plasmid R6K DNA replication II. Direct nucleotide sequence repeats are required for an active  $\gamma$ -origin. *J. Mol. Biol.* **161**, 45-56.
  33. Szybalski, W. 1993. From the double-helix to novel approaches to the sequencing of large genomes. *Gene* **135**, 279-290.
  34. Vetter, D., Andrews, B. J., Roberts-Beatty, L. and Sadowski, P. D. 1983. Site-specific recombination of yeast 2- $\mu$ m DNA *in vitro*. *Proc. Natl. Acad. Sci. USA* **80**, 7284-7288.
  35. Watson, J. D. 1990. The human genome project: past, present, and future. *Science* **248**, 44-49.
  36. Watson, J. D. and Crick, F. H. C. 1953a. A structure for deoxyribose nucleic acids. *Nature* **171**, 737-738.
  37. Wright, E. M., Humphreys, G. O. and Yarranton, G. T. 1986. Dual-origin plasmids containing an amplifiable *ColE1 ori*: temperature-controlled expression of cloned genes. *Gene* **49**, 311-321.
  38. Yates, J. L., Warren, N. and Sugden, B. 1985. Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells. *Nature* **313**, 812-815.
  39. Yarranton, G. T., Wright, E., Robinson, M. K. and Humphreys, G. O. 1984. Dual-origin plasmid vectors whose origin of replication is controlled by the coliphage lambda promoter *p<sub>L</sub>*. *Gene* **28**, 283-300.

### 스트레스 해소의 건강음료 “진한대추”

제일제당(대표 손경식)은 최근 옛날부터 한방과 민간에서 강장·완화제로 애용되어온 대추를 주원료로한 스트레스 대응 건강음료 “진한대추”를 출시했다. 대추가 위장기능을 개선하고 신경안정 등 진정작용에 효과가 있다는 점에 착안하여 제조된 “진한대추”는 대추엑기스 함량이 1만 5천mg으로서, 부드러운 대추의 맛을 강조한 스트레스 대응 건강음료로서 스트레스가 많은 직장인, 수험생 및 중년 여성층을 대상으로 출시하였다.